Maximilian Huber

+1 (443)-859-7702 | huber.maxi@northeastern.edu | max-huber888 | MaxHuber888 | Boston, MA

Education

Northeastern University, Khoury College of Computer Science — Boston, MA

M.S. in Artificial Intelligence · ***Current GPA: 3.83 / 4.00*** — *Sep. 2024 – Present*

B.S. in Computer Science, Concentration in Artificial Intelligence · ***Final GPA: 3.72 / 4.00*** — *Sep. 2020 – Aug. 2024*

- **University Honors Distinction:** Awarded to Northeastern University students who have completed six Honors courses or approved experiential learning experiences and maintained a cumulative 3.500 GPA.

Skills

Programming: Python, Java, C, SQL, JavaScript, DLang, HTML/CSS, React/React Native, TypeScript

Libraries/APIs: TensorFlow, PyTorch, OpenAI, Langchain/LangGraph, LlamaIndex, OpenCV, NLTK

Software: GitHub, Conda, Poetry, Vim, Redis, GCS(BigQuery, Firebase), Apache Airflow, Metabase

Work Experience

Blorm — Remote

AI/ML Engineer — *Nov. 2024 – Present*

- Initialized and developed ZerePy, an open-source Python framework that facilitates the creation and launching of autonomous LLM agents with social media and blockchain functionality. The repository has over 500 stars on GitHub and has been used as a base for multiple hackathons.
- Designed and implemented an agentic autonomous decision graph using LangGraph. The graph allows an AI agent to retrieve and observe context, autonomously determine complex tasks to perform, and orchestrate various actions in sequence to accomplish those tasks.

CollX — Haddonfield, NJ, United States

Full Stack Engineer [Co-op] — *Mar. 2023 – Dec. 2023*

- Designed and implemented a Retrieval-Augmented Generation (RAG) AI chatbot that answers user queries about trading cards. Achieved a 90% reduction in average query time and enhanced accuracy by integrating a continuously updated knowledge index that automatically ingests relevant articles and documents.
- Leveraged existing SQL data tables to create a KPI dashboard which was used to predict and monitor app growth and marketplace activity in Metabase. Insights were used to identify marketplace trends and future marketplace ambassador candidates.
- Created and scheduled Apache Airflow jobs for ETL pipelines, enhancing visual search accuracy and automating AI knowledge index updates.

Projects

DeepSquid (Grad AI Final Project) | *Recurrent Neural Network (RNN) for deepfake video detection*

- Designed, implemented, and trained an RNN model using TensorFlow and OpenCV for detecting whether a given video has been deepfaked. Model reached 100% validation accuracy during training, and outperformed Mesonet and Mouthnet in testing and training.
- Designed a web demo using Vercel and Gradio which allows for testing the model on public YouTube URLs.

Chunk Splitting Problem (NLP Final Project) | *Experimental approach to indexing a text corpus for RAG*

- Identified and formalized the ”Chunk Splitting Problem”, which negatively impacts performance of all popular indexing methods when indexing an unstructured text corpus.
- Designed, implemented, and tested an experimental indexing method, which uses a PyTorch neural network to learn a representation of a text corpus, which circumvents chunk splitting problem.
- Created a testing framework for implementing and comparing different indexing methods in terms of embedding cost, time taken, and quality of retrieval results.

MindReader Quantum (MIT iQuHack Project) | *Classical-to-Quantum CNN for diagnosing dementia severity*

- Researched, implemented, and trained a Classical-to-Quantum hybrid CNN model to detect the level of dementia severity from a patient’s MRI brain scans. The hybrid model uses a pre-trained classical neural network for feature extraction and a quantum neural network decoder for efficient parallelizable training.
- Connected the backend to a React Native frontend via a REST API, later switched to a Gradio interface.